Our File: 9-13528-61US MRW:kp

#8/B
L.McLaughlin
1.22.02

RECEIVED
JAN 2 2 2003
Technology Center 2600

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re Application of: Maged E. Beshai

Application No: 09/550,489

Filed: April 17, 2000

Subject: HIGH-CAPACITY WDM-TDM PACKET SWITCH

Group Art Unit: 2739

THE COMMISSIONER OF PATENTS AND TRADEMARKS,
WASHINGTON, D.C. 20231,
U.S.A.

FAX RECEIVED
JAN 2 1 2003
GROUP 1700

Sir:

PRELIMINARY AMENDMENT

Please amend the above-referenced application, as follows:

IN THE DISCLOSURE:

Please replace the paragraph beginning on page 2, line 25 and ending on page 3, line 5 with the new paragraph set forth below:

B1

This is accomplished in a rate-controlled multi-class high-capacity packet switch described in Applicant's copending United States Patent Application No. 09/244,824 which was filed on February 4, 1999. Although the switch described in this patent application is adapted to switch variable sized packets at very high speeds while providing grade-of-service and quality-of-service control, there still exists a need for a distributed switch that can form the core of a powerful high-capacity, high-performance network that is adapted to provide wide geographical coverage with end-to-end capacity that scales to hundreds of Tera bits per second (Tbs), while providing grade of service and quality of service controls.

Please replace the paragraphs beginning on page 4, line 31 and ending on page 7, line 27 with the new paragraphs set forth below:

A very high-capacity packet switch is adapted to provide a high service-quality, as well as providing intra-switch data paths with a fine granularity that reduces or eliminates a requirement for tandem switching. The packet switch requires a scheduler to coordinate the transfer of packets across the switch, and the scalability of the switch is primarily determined by the throughput of its scheduler. Providing a fine granularity in a high-capacity switch requires an extensive scheduling effort that may not be realizable with a single controller. The invention, therefore, provides a switch that includes a plurality of core modules that operate in a time-division mode, and a plurality of edge modules that are connected to subtending packet sources and sinks, with each core module having its own controller which includes a packet-transfer scheduler.

B2 cont

In accordance with an aspect of the present invention, there is provided a packet switch. The packet switch comprises a plurality of independently-controlled core modules, a plurality of ingress modules, and a plurality of egress modules. The packet switch may further include a plurality of core controllers operating concurrently and independently; one core controller associated with each of the independently-controlled core modules and having a packet scheduler. Each ingress module receives packets from subtending packet sources and has a link directed to each of the core modules. Each egress module has a link from each of the core modules and transmits packets to subtending packet sinks. Each of the ingress modules is operable to issue packet-transfer requests and distribute the packet-transfer requests among the core modules for scheduling. Each core module computes schedules in response to receiving packet-transfer requests, the schedules specifying time slots in a predefined time frame for each request. The ingress modules and the core modules can be geographically distributed and each ingress module is provided with a plurality of timing circuits each communicating with a time counter associated with one of the core modules to realize time coordination between each ingress module and the core modules.

In accordance with another aspect of the present invention, there is provided a method of scheduling. The method is performed by a controller of a core module having $S \geq 1$ space switches and at least one link to each of a

plurality of ingress modules, where each ingress module formulates capacity-allocation requests preferably organized in capacity-request vectors each entry of which specifying an input port p, an output port π, and a number K of time slots per time frame. The method relies on a data structure to facilitate the scheduling process. The data structure preferably comprises a first three-dimensional matrix having a space dimension s representing space switches associated with the core module, a space dimension p representing space-switch input ports, and a time dimension t representing the time slots in a slotted frame, and a second three-dimensional matrix having the space dimension s, a space dimension π representing space-switch output ports, and said time dimension t. The method comprises steps of creating the data structure, receiving capacity-allocation requests from the ingress edge modules, selecting a space switch s and a time slot t and, if both entries {s, p, t} of the first three-dimensional matrix and {s,π, t} of the second three-dimensional matrix are free, allocating the space switch s and the time slot t and marking entries {s, p, t} and {s,π, t} as busy. The step of selecting is repeated until at most K time slots are allocated. The method includes the further step of terminating a current connection by setting the value of K to equal to zero.

B2
Cont

In accordance with a further aspect of the present invention, there is provided a distributed packet switch. The distributed packet switch comprises a plurality of m cross connectors, a plurality of n core modules, a plurality of m×n edge modules, and a plurality of n core controllers each having a core scheduler. Each cross connector has n outer links and n inner links. Each outer link connects to an edge module and includes Λ channels in each direction to and from the edge module. Each inner link connects to a core module and includes Λ channels in each direction to and from the core module. Each core module comprises a number of space switches not exceeding the ratio Λ/n. The edge modules and the core modules can be spatially distributed over a wide geographical area and the outer and inner links are preferably wavelength-division-multiplexed links. Each edge module has means for time coordination with the core modules. The core

controller of any core module is adapted to compute a schedule in response to receiving capacity-allocation requests, the schedule specifying, for each capacity-allocation request, time slots in a predefined time frame.

In accordance with a still further aspect of the present invention, there is provided a packet switch. The packet switch comprises a plurality of egress modules, each for transmitting packets on at least one network link, a plurality of ingress modules, each for receiving packets from at least one network link and capable of requesting ingress-to-egress-module connections for transferring received packets to any other of the egress modules, and a plurality of core modules, each capable of simultaneously receiving and independently responding to the ingress-to-egress-module connection requests from any of the ingress modules and of providing the ingress-to-egress module connections between any of the ingress modules and any of the egress modules in response to the connection requests. Each core module may have its own controller for allocating and scheduling resources to the ingress-to-egress-module connections. A core controller operates independently of, and concurrently with, the other core modules' controllers. Each edge module has a plurality of ingress ports each having an associated ingress buffer for receiving packets from subtending packet sources. An ingress controller in each ingress module sorts packets arriving in the ingress buffer into ingress queues, each ingress queue corresponding to an egress module from which packets are to egress from the switch for delivery to subtending packet sinks. Each edge module has a number of timing circuits at least equal to the number of core modules, each of the timing circuits being time-coordinated with a time counter associated with each of the core modules.

B2 Cont.

In accordance with an additional aspect of the present invention, there is provided a method of switching packets through a switch having a plurality of ingress modules each having at least one ingress port, a plurality of egress modules each having at least one egress port, and a plurality of core modules. Each ingress module is coupled to each core module, each core module is coupled to each egress module, and a packet can traverse only one ingress module, one core module, and one egress module in moving from an ingress port to an egress port. The method comprises steps of receiving, at an ingress

module, packets from subtending traffic sources, the ingress module selecting the egress modules to which to send the packets, sending connection requests to selected core modules, and requesting connections of specified capacities. The method includes the further steps of determining a feasible capacity allocation in response to a connection request, subtracting the feasible capacity allocation from a specified capacity, and returning an updated connection request to the ingress module that issued the connection request. If the feasible capacity allocation is less than the specified capacity, the ingress module may send the connection request to another core module.

B2 Contd.

In accordance with another aspect of the present invention, there is provided an ingress module in a packet switch. The ingress module comprises an ingress controller, a plurality of ingress ports each having an ingress buffer for receiving packets from subtending packet sources where each packet indicates one of predefined destinations, a plurality of output ports for directing the packets to a plurality of core modules, means for sorting the packets received in the ingress buffer into ingress queues each corresponding to one of the destinations, means for storing a set of predefined paths to each of the predefined destinations, means for formulating connection requests, each connection request specifying a destination and a required capacity allocation, and means for selecting a candidate path from among the predefined paths for each connection request.

In accordance with a further aspect of the present invention, there is provided a core module in a packet switch. The core module comprises at least one space switch having a plurality of input ports and a plurality of output ports, and a core controller adapted to receive connection requests, each connection request specifying a required capacity allocation and a destination selected from among a set of predefined destinations. The core controller provides means for associating each destination with one of the output ports, and a scheduler associated with the core controller times the transfer of packets from the input ports to the output port and communicates scheduling results to sources of the connection requests.

Please replace the paragraphs that begin on page 21, line 27 and ends on page 22, line 32 with the amended paragraphs set forth below:

The capacity-request matrix 44 sent to a core module 34 is normally a sparse matrix with a majority of null entries since the capacity demand is split among eight core modules. The controller for a core module attempts to schedule the capacity requested by each ingress edge module 22 using data structures generally indicated by references 46 and 48. Each of the data structures 46, 48 is a three-dimensional matrix having a first space dimension *s*, which represents the respective space switches associated with the core module 34; a second space dimension *p*, which represents the space-switch ports; and a time dimension *t*, which represents the slots in a slotted frame. Thus, an entry in data structure 46 is represented as *{s,p,t}*. The second dimension *p* may represent an input channel, if associated with the data structure 46, or an output channel if associated with the data structure 48. If the number of slots *T* per frame is 16, for example, then in the configuration of FIG. 1, which shows a centralized core, the size of the three-dimensional structure 46 is $128 \times 256 \times 16$. In the distributed core shown in FIG. 3, each core module uses a three-dimensional structure 46 of size $16 \times 256 \times 16$.

33

When the connections through a core module 34 are reconfigured, the core controller may either reschedule the entire capacity of the respective core module 34 or schedule capacity changes by simply perturbing a current schedule. If the entire capacity of the core module is reconfigured, each ingress edge module 22 must communicate a complete capacity request vector to the core module while, in the latter case, each ingress edge module 22 need only report capacity request changes, whether positive or negative, to a respective core controller. A negative change represents capacity release while a positive change indicates a request for additional capacity. The incremental change method reduces the number of steps required to prepare for reconfiguration.

Please replace the paragraph that begins on page 23, line 21 and ends on page 24, line 2 with the amended paragraph set forth below: